Created by Neil Yetz & Gemma Wallace

Module 5 Lab Activity: Data Wrangling Practice
PSY 652 Research Methods
Oct 2, 2019

```
---
title: "Fandango_Notebook"
subtitle: "PSY 652 RM Module 5 Lab"
output:
  html_notebook:
    toc: yes
---
```

This activity is designed to build familiarity with data wrangling techniques in R. In this activity you will create a new R notebook using the Fandango dataset. This is the raw data behind the story "Be Suspicious Of Online Movie Ratings, Especially Fandango's. (http://fivethirtyeight.com/features/fandango-movies-ratings/)." The dataframe contains every film that has a Rotten Tomatoes rating, a RT User rating, a Metacritic score, a Metacritic User score, and IMDb score, and at least 30 fan reviews on Fandango.

```
# Load libraries
```{r, message = FALSE}
library(tidyverse)
```
```

```
# Import data
```{r, message = FALSE}
fandango <- read_csv("fandango_module5lab.csv")
```
```

```
# See the structure of the data frame
```{r}
str(fandango)
```
```

```
# Examine descriptives using the summary function
```{r}
summary(fandango)
```
```

A collection of data wrangling techniques

filter the data
The filter function allows you subset data by rows.
```{r}
film_2014 <- filter(fandango, year==2014)
film_2014 <- filter(fandango, year==2014 & stars >=4)
```

select the data
The select function allows you to subset data by column.
```{r}
fandango_subset <- select(fandango, year, stars)
```

mutate the data
The mutate function allows you to create new variables in your dataset. The mutate function can either create new variables or modify existing variables.
```{r}
fandango <- mutate(fandango, stars_ten = stars*10)
```

arrange the data
The arrange function allows you to sort the data by certain variable(s).
```{r}
fandango_sorted <- arrange(fandango, year)
```

summarise the data
The summarize function allows you to collapse cases for easier descriptive interpretations and plotting.
```{r}
fandango_year <- group_by(fandango,year)
summarise(fandango_year, year_mean = mean(stars))
```